PEER REVIEW HISTORY

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Safety and efficacy of hydroxychloroquine as prophylactic against COVID-19 in healthcare workers: a meta-analysis of randomized clinical trials
AUTHORS	Hong, Hwanhee; Friedland, Anne; Hu, Mengyi; Anstrom, Kevin J.; Halabi, Susan; McKinnon, John; Amaravadi, Ravi; Rojas-Serrano, Jorge; Abella, Benjamin; Portillo-Vázquez, Angélica Margarita; Woods, Christopher; Hernandez, Adrian F.; Boulware, David R.; Naggie, Susanna; Rajasingham, Radha

VERSION 1 – REVIEW

REVIEWER	Fiolet, Thibault Université Paris-Saclay, Espace Maurice-Tubiana
REVIEW RETURNED	13-Jul-2022

GENERAL COMMENTS	This meta-analysis assessed the impact of hydroxychloroquine use as prophylaxis against suspected+confirmed or confirmed SARS-CoV-2 infection among healthcare workers (HCW). They identified 5 randomized trials and the pooled analysis shows that there is not effect of HCQ on SARS-CoV-2 infection nor on adverse events. The manuscript is clear and the results are useful for clinicians and policy makers. Methodology of systematic review and PRISMA checklist are followed. I agree with the authors that it may not be possible to conduct subgroup analyses due to the lack of information and the inclusion of only 5 trials. I have few comments and questions: 1) I wonder if the systematic review should be updated. Last search was conducted on October 11, 2021. Rojas Serrano has been published in PLos One. I found new trials which may be added: https://bmjopen.bmj.com/content/12/6/e059540.abstract https://www.medrxiv.org/content/10.1101/2022.03.02.22271710v1 2) The statistical method sounds appropriate but were the Odds Ratio adjusted? Or was it just calculated using the number of cases/non-cases in each arms? 3) What is the advantage of using a Bayesian approach if no priori information was available? 4) Table 1: is it only included studies? It may be specified in the title please. Please could you add the authors names in addition to the names of the trials?

	5) I wonder if the search terms are not too restrictive (it was limited to titles and abstract)

REVIEWER	Carayannopoulos, Kallirroi Laiya McMaster University
REVIEW RETURNED	12-Oct-2022

GENERAL COMMENTS	Overall a well written and statistically robust SRMA. I have two minor concerns: 1) While the authors do briefly reference that vaccine availability remains limited on a global level and thus HCQ continues to be worth exploring as prophylaxis, they do not elaborate on vaccination rates of healthcare workers specifically. Does this same disparity in vaccine access apply to HCWs around the globe? A brief discussion on why prophylaxis in the era of vaccines would be valuable in this specific population would add strength to the relevance of this review. 2) The PRISMA checklist states that information regarding the certainty of evidence is available in the supplement, however, I cannot find this information listed there, nor discussed throughout the body of the text. Was the GRADE approach applied to these outcomes and simply not discussed? If not, I would recommend completing an assessment of all outcomes and making this information available as well as discussing it in the context of the findings.

REVIEWER	Horvath, Laszlo University of Exeter
REVIEW RETURNED	29-Dec-2022

GENERAL COMMENTS	Statistical Reviewer - the authors present a Bayesian random effects meta analysis, which they describe very well, and provide a thorough discussion of the results. I have two small notes: - it might be helpful for some readers if the authors added a sentence and/or pointed to a reference to explain why the Bayesian framework is preferable here over a frequentist RE model, given the uninformative priors used in the study. Is this choice perhaps informed by the sample size? New practices, new conventions in the field? - for readers to get an idea of the research landscape, it might be helpful to explain what the database search results represent. One scenario is that we can interpret the 3/164 split as the HCQ research is predominantly observational, with only 3 RCTs. But this doesn't seem right because the query given eTable1 suggests RCT was part of the keywords. So what exactly do these numbers represent in non-technical terms? Perhaps interpret this in Methods/Eligibility criteria and study selection.

REVIEWER	Mehta, Kedar GMERS Medical College Gotri Vadodara, Community Medicine Department
REVIEW RETURNED	19-Jan-2023

GENERAL COMMENTS	Congratulations to the team for good systematic work done for addressing the research question. However, I have some major concerns with the research topic. 1. The topic is well known to the scientific forum. At present, HCQ among health care workers is not recommended. So, this research question is well answered by earlier studies and meta-analysis.

	There is nothing novel in the manuscript. Kindly justify the rationale in conducting this study. 2. Sensitivity analysis for each outcome variable is not conducted. Also, it would be good if funnel plots can be added as supplementary files to indicate publication bias. 3. Figure 1 - PRISMA flow diagram shows that 3 unpublished trials have been included - kindly explain reasons for the same for including unpublished work in the meta-analysis - how those unpublished work has affected results/outcome (if any)? 4. Geographical variation needs to be discussed in discussion section.

REVIEWER	Khamis, Assem Hull York Medical School, Wolfson Palliative Care Research Centre
REVIEW RETURNED	12-Feb-2023

GENERAL COMMENTS	I have a minor revision in the abstract (page 4, line 84) and results (page 12, line 249): I wonder if the authors wanted to report credible intervals or confidence intervals because there is inconsistency across the manuscript. I have no further comments.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 (Dr. Thibault Fiolet, Université Paris-Saclay)

This meta-analysis assessed the impact of hydroxychloroquine use as prophylaxis against suspected+confirmed or confirmed SARS-CoV-2 infection among healthcare workers (HCW). They identified 5 randomized trials and the pooled analysis shows that there is not effect of HCQ on SARS-CoV-2 infection nor on adverse events. The manuscript is clear and the results are useful for clinicians and policy makers. Methodology of systematic review and PRISMA checklist are followed. I agree with the authors that it may not be possible to conduct subgroup analyses due to the lack of information and the inclusion of only 5 trials.

I have few comments and questions:

C2: I wonder if the systematic review should be updated. Last search was conducted on October 11, 2021. Rojas Serrano has been published in PLos One. I found new trials which may be added:
https://urldefense.com/v3/__https://bmjopen.bmj.com/content/12/6/e059540.abstract__;!!OToaGQ!oS3n2iWMxQVDXPk_av34nONwTKfBHBEtW2PZbC_kG_Ie8sNHrp168BD-AoPMtY6BN_-nywNnzmvE8aYdtqTccBpyddzqzjXd$
https://urldefense.com/v3/__https://www.medrxiv.org/content/10.1101/2022.03.02.22271710v1__;!!OToaGQ!oS3n2iWMxQVDXPk_av34nONwTKfBHBEtW2PZbC_kG_Ie8sNHrp168BD-AoPMtY6BN_-nywNnzmvE8aYdtqTccBpyda3dXUOr$

R2: We thank the reviewer for this comment and we agree that the systematic review became a bit outdated because the manuscript was submitted for review about 9 month ago. We have now updated the systematic review as of March 14, 2023. We have identified 5 additional eligible trials, resulting in a total of 10 trials included in the meta-analysis. Note that the 5 additional trials were negative trials concluding that hydroxychloroquine is not effective to prevent COVID-19 infection compared with placebo. As such adding these 5 trials to our meta-analysis strengthened the original conclusion. A summary table comparing results of the original and updated meta-analyses is provided below. The conclusions about the two efficacy outcomes (lab-confirmed positive COVID-19 and

suspected COVID-19) remain the same, while the odds ratio for adverse events became statistically marginally significant, supporting that HCQ is less safe than placebo.

	Odds ratio (HCQ vs. placebo) in the original meta-analysis (# of studies=5)	Odds ratio (HCQ vs. placebo) in the updated meta-analysis (# of studies=10)
Lab-confirmed positive COVID-19	0.60 (0.24, 1.28)	0.92 (0.58, 1.37)
Suspected COVID-19	0.76 (0.48, 1.24)	0.78 (0.57, 1.10)
Adverse events	1.46 (0.87, 2.22)	1.35 (1.03, 1.73)

We have updated reference for the Rojas Serrano study using the publication in PLoS One. Additionally, we have updated reference for the HERO-HCQ study.

C3: The statistical method sounds appropriate but were the Odds Ratio adjusted? Or was it just calculated using the number of cases/non-cases in each arms?

R3: The odds ratios were not adjusted because we do not have individual participants-level data to estimate the adjusted odds ratios. We pooled the number of cases/non-cases reported individual studies using a Bayesian logistic regression meta-analysis model with random effects to account for the variability in odds ratios across trials.

C4: What is the advantage of using a Bayesian approach if no priori information was available?

R4: The Bayesian meta-analysis approach has several advantages. First, its flexibility and the MCMC sampling methods to estimate posterior distributions provide probability-based quantities (e.g., posterior probability of an odds ratio smaller than 0.5) that complement typical meta-analysis results (e.g., odds ratios and the associated credible intervals) and help decision making [1]. Second, the Bayesian meta-analysis model with random effects estimates the between-study variability better than the frequentist counterparts [2]. Third, when it comes to with binary outcomes, the Bayesian approach handles rare events better than the frequentist counterparts [3]. We have now added these points in the Discussion section.

[1] Hong H, Chu H, Zhang J, Carlin BP. A Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons. Research synthesis methods. 2016 Mar;7(1):6-22.

[2] Hong H, Carlin BP, Shamliyan TA, Wyman JF, Ramakrishnan R, Sainfort F, Kane RL. Comparing Bayesian and frequentist approaches for multiple outcome mixed treatment comparisons. Medical Decision Making. 2013 Jul;33(5):702-14.

[3] Hong H, Wang C, Rosner GL. Meta-analysis of rare adverse events in randomized clinical trials: Bayesian and frequentist methods. Clinical Trials. 2021 Feb;18(1):3-16.

C5: Table 1: is it only included studies? It may be specified in the title please. Please could you add the authors names in addition to the names of the trials?

R5: Thank you for the catch. We have now modified the caption for Table 1 to specify these are for trials included in the meta-analysis. Also, we added authors names and citations in the table body.

C6: I wonder if the search terms are not too restrictive (it was limited to titles and abstract)

R6: One of the most important eligibility criteria is randomized controlled trials (RCT). Most clinical journals require that articles about RCTs specify their study designs in the title and abstract clearly.

As such, titles and abstract search was sufficient to identify relevant RCTs. While we updated our systematic review to respond C2, we revisited the search terms. Surprisingly, using a small set of search keywords limited to titles only identified more relevant papers than using the same set of search keyword limited to titles and abstract. As a result, the former was used in the updated systematic review. Note that we dropped the EBSCO and Cochrane databases completely from the search because we lost access to EBSCO (no longer available from the first author's institution library) and the Cochrane database provided a very few irrelevant papers.

Reviewer: 2 (Dr. Kallirroi Laiya Carayannopoulos, McMaster University)
Comments to the Author:
Overall a well written and statistically robust SRMA. I have two minor concerns:

C7: While the authors do briefly reference that vaccine availability remains limited on a global level and thus HCQ continues to be worth exploring as prophylaxis, they do not elaborate on vaccination rates of healthcare workers specifically. Does this same disparity in vaccine access apply to HCWs around the globe? A brief discussion on why prophylaxis in the era of vaccines would be valuable in this specific population would add strength to the relevance of this review.

R7: Thank you for this comment. We have added additional details regarding the proportion of healthcare workers vaccinated globally in the Introduction section. Specifically, in low-income countries only 33% of healthcare workers are fully vaccinated. While high-income countries have better coverage, overall 38% of countries did not achieve the milestone of 70% vaccination coverage for healthcare workers by the end of 2021. [4]

[4] Nabaggala MS, Nair TS, Gacic-Dobo M, Siyam A, Diallo K, Boniol M. The global inequity in COVID-19 vaccination coverage among health and care workers. International Journal for Equity in Health. 2022 Oct 13;21(Suppl 3):147.

C8: The PRISMA checklist states that information regarding the certainty of evidence is available in the supplement, however, I cannot find this information listed there, nor discussed throughout the body of the text. Was the GRADE approach applied to these outcomes and simply not discussed? If not, I would recommend completing an assessment of all outcomes and making this information available as well as discussing it in the context of the findings.

R8: Thank you for pointing this out. We have now mentioned the GRADE approach to the newly named subsection "Risk of bias and certainty of evidence assessment," and the associated results in the Results of meta-analysis subsection. The GRADE score for the odds ratios was downgraded by 1 because a few studies showed wide credible intervals of odds ratios, resulting in moderate certainty of evidence for all outcomes.

Reviewer: 3 (Dr. Laszlo Horvath, University of Exeter)
Comments to the Author:
Statistical Reviewer - the authors present a Bayesian random effects meta analysis, which they describe very well, and provide a thorough discussion of the results. I have two small notes:

C9: it might be helpful for some readers if the authors added a sentence and/or pointed to a reference to explain why the Bayesian framework is preferable here over a frequentist RE model, given the uninformative priors used in the study. Is this choice perhaps informed by the sample size? New practices, new conventions in the field?

R9: Thank you for your comment. Reviewer 1 provided a similar comment (please see R4 for the response). We have now added these points in the Discussion section.

C10: for readers to get an idea of the research landscape, it might be helpful to explain what the database search results represent. One scenario is that we can interpret the 3/164 split as the HCQ research is predominantly observational, with only 3 RCTs. But this doesn't seem right because the query given eTable1 suggests RCT was part of the keywords. So what exactly do these numbers represent in non-technical terms? Perhaps interpret this in Methods/Eligibility criteria and study selection.

R10: Our search results depend on the keyword search algorithms provided by PubMed, EMBASE, EBSCO, and Cochrane. What we learned during this search is that the keyword search provided a generously inclusive list of articles. For example, if an article mentioned some keywords in texts (e.g., Introduction), it could be selected. For this reason, most systematic reviews start with a large number of articles. This is why it is crucial to predetermine eligibility criteria and apply them accurately to screen a large number of articles. As we answered in R6, we updated our systematic review in this revision per Reviewer 1's comment. While we updated our systematic review, we revisited the search terms. Surprisingly, using a small set of search key word limited to titles only identified more relevant papers than using the same set of search key word limited to titles and abstract. As a result, the former was used in the updated systematic review.

Reviewer: 4 (Dr. Kedar Mehta, GMERS Medical College Gotri Vadodara)
Comments to the Author:
Congratulations to the team for good systematic work done for addressing the research question. However, I have some major concerns with the research topic.

C11: The topic is well known to the scientific forum. At present, HCQ among health care workers is not recommended. So, this research question is well answered by earlier studies and meta-analysis. There is nothing novel in the manuscript. Kindly justify the rationale in conducting this study.

R11: Thank you for this comment. Multiple trials have been conducted, many of which were underpowered. We think there is value in combining these data for further analysis. Furthermore, we believe it is important to publish and analyze null data to contribute to the scientific literature. There are many who still believe hydroxychloroquine has a role in preventing Covid-19 infection. We are trying to combine the existing literature, even if studies are underpowered, to definitively answer this question. We have now mentioned this in the last paragraph of the Discussion section.

C12: Sensitivity analysis for each outcome variable is not conducted. Also, it would be good if funnel plots can be added as supplementary files to indicate publication bias.

R12: We thank the reviewer for bringing this point. However, we did not plan to conduct sensitivity analysis because our updated meta-analysis included a fairly small number of RCTs (n=10) and the ten RCTs have similar population (healthcare workers at risk). As we mentioned in the Discussion section, subgroup analyses were not conducted due to limited information that can be obtained from published papers.

We also thank the reviewer for the suggestion about funnel plots to assess publication bias. We have now added the fennel plots in the supplemental document (eFigure 1) and additionally conducted the Egger's test (results shown in the Results of meta-analysis subsection). Funnel plots and Egger's test showed no indication of publication bias.

C13: Figure 1 - PRISMA flow diagram shows that 3 unpublished trials have been included - kindly explain reasons for the same for including unpublished work in the meta-analysis - how those unpublished work has affected results/outcome (if any)?

R13: Per Reviewer 1's comment, we have updated our systematic review as of March 14th, 2023 (see our response R2 above) and meta-analysis accordingly. The three unpublished studies at our manuscription submission were now published in peer-reviewed journals. We have now updated the flow diagram (Figure 1) and reference accordingly.

C14: Geographical variation needs to be discussed in discussion section.

R14: Our updated meta-analysis included a total of 10 trials. The geographical locations of the 10 trials included in the meta-analysis are US, Canada, Mexico, India, Spain, Bolivia, Venezuela, Peru, and Pakistan (eTable 3). While the odds ratios of most studies favor HCQ, the credible intervals remain wide suggesting low certainty in the true point estimate. Two studies including the Llanos-Cuentas et al. study conducted in Peru and the Syed et al. study conducted in Pakistan showed odds ratios favoring placebo, though the credible intervals remain wide.

Reviewer: 5 (Dr. Assem Khamis, Hull York Medical School)
Comments to the Author:

C15: I have a minor revision in the abstract (page 4, line 84) and results (page 12, line 249): I wonder if the authors wanted to report credible intervals or confidence intervals because there is inconsistency across the manuscript. I have no further comments.

R15: Thank you for catching it! We meant credible intervals because all results were based on Bayesian models. We corrected these typos and now they all read credible intervals consistently throughout the manuscript.

VERSION 2 – REVIEW

REVIEWER	Carayannopoulos, Kallirroi Laiya McMaster University
REVIEW RETURNED	21-Apr-2023

GENERAL COMMENTS	I thank the authors for their revisions and updated search of available literature. I do note the clarification regarding assessment of the certainty of evidence, however, I do have some ongoing concerns with the reporting of results. Specifically, there is now a reference to each of the outcome data being downgraded to moderate certainty due to the confidence intervals, however, this is not reflected in the presentation of the data. For example, adverse events are described as having "marginally statistical significance." This language is meaningless as there are not degrees of statistical significance and renders the purpose of assessing the certainty of evidence moot. Having performed the certainty assessment, statements of results should read something to the effect of "Participants treated with HCQ had a higher rate of adverse events (OR 1.35, 95% CI: 1.03, 1.73; moderate certainty)". Please revise the manuscript to appropriately integrate the certainty of evidence and remove any inappropriate qualifiers regarding statistical significance. Finally, I have not been able to find the Summary of Findings table for the certainty assessment. Please upload and include it in the supplementary materials with any resubmission.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2 (Dr. Kallirroi Laiya Carayannopoulos, McMaster University)
C2: I thank the authors for their revisions and updated search of available literature. I do note the clarification regarding assessment of the certainty of evidence, however, I do have some ongoing concerns with the reporting of results.
Specifically, there is now a reference to each of the outcome data being downgraded to moderate certainty due to the confidence intervals, however, this is not reflected in the presentation of the data. For example, adverse events are described as having "marginally statistical significance." This language is meaningless as there are not degrees of statistical significance and renders the purpose of assessing the certainty of evidence moot. Having performed the certainty assessment, statements of results should read something to the effect of "Participants treated with HCQ had a higher rate of adverse events (OR 1.35, 95% CI: 1.03, 1.73; moderate certainty)". Please revise the manuscript to appropriately integrate the certainty of evidence and remove any inappropriate qualifiers regarding statistical significance. Finally, I have not been able to find the Summary of Findings table for the certainty assessment. Please upload and include it in the supplementary materials with any resubmission.

R2: Thank you for your comment and suggestion regarding how to state certainty assessments. We have made the suggested modifications to the Results of meta-analysis subsection and removed the word “marginally” when stating the adverse event results. For example, we now stated that “Participants treated with HCQ had a numerically higher rate of adverse events (OR 1.35, 95% CI: 1.03, 1.73; GRADE score: moderate certainty) with statistical significance.” In addition, we have included the Summary of GRADE score assessment in eTable 7.

VERSION 3 – REVIEW

REVIEWER	Carayannopoulos, Kallirroi Laiya McMaster University
REVIEW RETURNED	25-May-2023

GENERAL COMMENTS	I thank the authors for their response to the feedback provided in the previous review, however, my concerns expressed at that time remain inadequately addressed. I do note in your response the language regarding statistical significance was addressed, however, the language remains unchanged in the manuscript in both the abstract (page 4 line 85 - and marginally significant difference in adverse events) and main text (page 13 line 253-254 - Participants treated with HCQ had a numerically higher rate of adverse events ... with marginally statistical significance). Please amend. With respect to the summary of findings table, the GRADE handbook provides detailed instruction on the requirements for the table, and can be found in section 4.3. The GradePro tool (https://gdt.gradepro.org/app/) is valuable in facilitating inclusion of all these requirements. Please complete the table to match the requirements.

VERSION 3 – AUTHOR RESPONSE

Reviewer 2 (Dr. Kallirroi Laiya Carayannopoulos, McMaster University)

C1: I thank the authors for their response to the feedback provided in the previous review, however, my

concerns expressed at that time remain inadequately addressed. I do note in your response the language

regarding statistical significance was addressed, however, the language remains unchanged in the manuscript in

both the abstract (page 4 line 85 - and marginally significant difference in adverse events) and main text (page

13 line 253-254 - Participants treated with HCQ had a numerically higher rate of adverse events ... with

marginally statistical significance). Please amend.

With respect to the summary of findings table, the GRADE handbook provides detailed instruction on the

requirements for the table, and can be found in section 4.3. The GradePro tool

(https://urldefense.com/v3/__https://gdt.gradepro.org/app/__;!!OToaGQ!r80OhTuPoJk1Hfr0-

JzaSftgQMn2sH9KdlELlX7R-Z-a--IYvyTlNXAsknxnkDlCewLR8plBnrM6zbIwfriO3vmn8MVvDRtc$) is

valuable in facilitating inclusion of all these requirements. Please complete the table to match the requirements.

R1: We apologize for overlooking other parts of the manuscript and not addressing the language issues of

marginal significance. We have now amended it throughout the manuscript.

Thank you for your suggestion regarding the summary of findings table. We have updated eTable 7 and

followed Example 2, as provided in GRADE Working Group grades of evidence

(https://gdt.gradepro.org/app/handbook/handbook.html). The revised eTable 7 includes the required

elements and provides detailed reasons for the GRADE decision.

eTable 7. GRADE summary of findings table

Outcomes	No of participants (studies) Follow up	Quality of the evidence (GRADE)	Odds ratio (95% Confidence Interval)
Lab-confirmed positive COVID-19	5039 (10 studies) From 28 days to 180 days	⨁⨁⨁⊖ Moderate1 due to imprecision	0.92 (0.58, 1.37)
Suspected COVID-19	4087 (5 studies) From 56 days to 180 days	⨁⨁⨁⊖ Moderate1 due to imprecision	0.78 (0.57, 1.10)
Adverse events	4979 (9 studies) From 56 days to 180 days	⨁⨁⨁⊖ Moderate2 due to imprecision	1.35 (1.03, 1.73)

95% confidence interval includes effect suggesting benefit as well as no benefit. 2 Although the 95% confidence interval includes an effect suggesting no benefit, we decided to downgrade it by one level because the lower limit is close to the null.

VERSION 4 – REVIEW

REVIEWER	Carayannopoulos, Kallirroi Laiya McMaster University
REVIEW RETURNED	30-May-2023

GENERAL COMMENTS	N/A